Phonology **25** (2008) 129–164. © 2008 Cambridge University Press
doi:10.1017/S0952675708001425 Printed in the United Kingdom

*On the perceptual origin of loanword adaptations: experimental evidence from Japanese**

Sharon Peperkamp
Inga Vendelin
Laboratoire de Sciences Cognitives et Psycholinguistique (DEC-ENS, EHESS, CNRS)

Kimihiro Nakamura
University of Tokyo

Japanese shows an asymmetry in the treatment of word-final [n] in loanwords from English and French: while it is adapted as a moraic nasal consonant in loanwords from English, it is adapted with a following epenthetic vowel in loanwords from French. We provide experimental evidence that this asymmetry is due to phonetic differences in the realisation of word-final [n] in English and French, and, consequently, to the way in which English and French word-final [n] are perceived by native speakers of Japanese. Specifically, French but not English word-final [n] has a strong vocalic release that Japanese listeners perceive as their native vowel [ɯ]. We propose a psycholinguistic model in which most loanword adaptations originate in perceptual assimilation, a process which takes place during perception and which maps non-native sounds and sound structures onto the phonetically closest native ones. We compare our model to alternatives couched within phonological theory.

1 Introduction

This article is concerned with the treatment of word-final [n] in English and French loanwords in Japanese. Consider the data in (1), showing that loanwords from English and French are not treated alike: loanwords from

* Research for this article was funded by a doctoral grant from the French ministry of Foreign Affairs and by a grant from the *Agence Nationale pour la Recherche* (05-BLAN-0065-01). For comments and discussion we would like to thank participants at the 15th International Congress of Phonetic Sciences, the Paris Workshop on Problems with Surface-Based Generalizations, the 12th Manchester Phonology Meeting, audiences at the University of Pennsylvania, the State University of New York at Stony Brook and Northwestern University, as well as Abigail Cohn, Emmanuel Dupoux, Larry Hyman, Gregory Iverson, Janet Pierrehumbert, Kie Zuraw and two anonymous reviewers.

English are always adapted with a final moraic nasal consonant, whereas loanwords from French invariably end in an epenthetic vowel.[1]

(1) a. *Loanwords from English*

pe**n**	<	pen		
wōkuma**n**	<	walkman		
monsū**n**	<	monsoon		

b. *Loanwords from French*

ka**nnu**	<	Cannes	[kan]	
parije**nnu**	<	parisienne	[paʁizjɛn]	'Parisian-FEM'
terī**nu**	<	terrine	[tɛʁin]	'pâté, terrine'

Similar data can be found in Shinohara (1997), who collected online adaptations of English and French words (as opposed to loanwords that are integrated in the Japanese lexicon). These adaptations were provided by Japanese speakers with a good knowledge of English and French respectively. Examples are shown in (2).

(2) a. *Online adaptations of English words*

pei**n**	<	pain		
orizi**n**	<	origin		
sarūteisjo**n**	<	salutation		

b. *Online adaptations of French words*

dua**nnu**	<	douane	[dwan]	'customs'
pishi**nnu**	<	piscine	[pisin]	'swimming pool'
purohe**nnu**	<	prochaine	[pʁɔʃɛn]	'next-FEM'

In standard accounts of loanword adaptations, the input to loanword adaptations is constituted by the surface form of the source language, and the adaptations are computed by the phonological grammar of the borrowing language (see, among many others, Hyman 1970, Jacobs & Gussenhoven 2000). In such accounts, different borrowing languages can treat a given surface form differently, due to the fact that their phonological grammars are not the same. However, the Japanese data seem to be problematic. Here, the same surface segment, [n], gives rise to different adaptations within the *same* borrowing language.

How can there be two different adaptations in Japanese for what superficially looks like the same input segment (in the sense that the native intuition of both English and French speakers is that the words in the examples above end in [n])? There are several possible answers to this puzzle. First of all, loanword adaptations might be influenced by the spelling of the source words. Under this account, the asymmetry would be due to differences in the English and French spellings of words ending in

[1] Throughout this article, we transcribe Japanese words by means of the *rōmaji* alphabet, which comes fairly close to a phonetic transcription. Important for the purposes of the present paper is that the grapheme<u>stands for the back unrounded vowel [ɯ], and that syllable-final <n> has various phonetic realisations, to be discussed in §2.

[n]. Hence this solution would state that the input to the adaptations consists of an orthographic representation, which is different in English and French. Moreover, the adaptations would not be computed by the phonological grammar but by the reading system of Japanese. Second, loanword adaptations might be calculated on the basis of the underlying rather than the surface form. English and French word-final [n], then, would be argued to derive from different underlying forms. The assumption is thus made that loanword adaptations are calculated by highly proficient bilinguals. The same assumption underlies the third possible solution, according to which global differences in the English and French phonologies are responsible for the observed asymmetry in the Japanese adaptations. It would thus be the adapters' knowledge of the phonological systems of the source languages, English and French, which influences the adaptations, and loanword adaptations would not be (solely) computed by the phonological grammar of the borrowing language, Japanese. Finally, loanword adaptations might be calculated on the basis of a fine-grained phonetic representation rather than the phonological surface representation in the source language. In other words, subtle differences in the phonetic realisation of word-final [n] in English and French would be responsible for the different adaptations in Japanese.

We can thus distinguish three different types of explanation, one orthographic, one phonological and one phonetic. In this article, we argue in favour of the third type of explanation. Our proposal is couched within the framework of Peperkamp & Dupoux (2003) and Peperkamp (2005), in which loanword adaptations are considered to be the formal reflex of perceptual assimilation, a process that applies during speech perception and that maps non-native sound structures onto the *phonetically* closest native ones. This process being computed by an acoustic distance metric, we depart from the idea that loanword adaptations are computed by the phonological grammar of the borrowing language. Rather, they are *influenced* by it, in that it is this grammar that determines which sounds and sound structures are available for non-native ones to be mapped onto.

We investigate the perceptual origin of the Japanese loanword data (and hence their phonetic nature) with two perception experiments. The first experiment shows that monolingual Japanese speakers perceive English word-final [n] to be closest to the moraic nasal consonant in their language and French word-final [n] to a sequence of nasal consonant plus vowel. The second experiment shows that bilingual Japanese speakers have difficulties distinguishing French forms ending in [n] from corresponding ones ending in [nu]. Together, these experiments show that Japanese speakers, firstly, are sensitive to fine phonetic differences between English and French word-final [n], and, secondly, perceive a vowel at the end of word-final [n] in French. Hence the loanword data are mirrored by the way in which Japanese listeners perceive word-final English and French [n]. Given the primacy of perception over production, this is strong evidence that the loanword adaptations originate in perceptual assimilation and consist of phonetically minimal transformations. We conclude by

comparing our psycholinguistic model to the linguistic model proposed by Kang (2003) and Kenstowicz & Suchato (2006) to account for the role of perception in loanword adaptations. Despite the presence of similarities with this model, our psycholinguistic model will be shown to be fundamentally different in one respect: due to the automatic character of perceptual assimilation and the primacy of perception over production, it only allows for a limited number of loanword adaptations that are *not* due to distortions during speech perception, but rather to adjustments taking place during production.

2 Absence *vs*. presence of epenthesis: orthography, phonology or phonetics?

Japanese has a very simple syllable structure. Branching onsets are not allowed, and syllable codas can be filled only by a moraic nasal consonant or by the first half of a geminate. In loanwords, input structures that violate the Japanese syllable structure constraints are typically repaired by means of vowel epenthesis (Lovins 1975). Whether or not the presence of a word-final [n] in foreign words violates Japanese phonotactics is debatable. This is due to the fact that the Japanese moraic nasal consonant is quite unlike [n] in most other languages, including English and French. Indeed, its phonetic realisation varies considerably, according to the phonological context. Before a consonant, it is realised as a nasal consonant that assimilates in place to the following consonant; in other contexts, it is realised as an unreleased uvular nasal, or even as a nasalised copy of the preceding vowel (Shibatani 1990, Kondo 1997). It is therefore not obvious whether English and French [n] can be said to correspond phonologically to the moraic nasal found syllable-finally in Japanese. For the present purposes, however, we do not need to argue one way or the other, since whichever position we take, the asymmetry between the absence of epenthesis in loanwords from English and its presence in loanwords from French remains surprising. On the one hand, if we consider final [n] in English and French words to be in accordance with the Japanese phonotactics, then it is the pattern with epenthesis in loanwords from French that seems unnecessarily complex. On the other hand, if we consider word-final [n] to constitute a violation of the Japanese phonotactics, then it is the English pattern that is deviant, in that the violation is not repaired by means of epenthesis, which is the usual repair strategy in case of syllable-structure violations.

As mentioned above, there are three possible types of account of the asymmetry, based on orthographic, phonological and phonetic arguments respectively. We will consider them in turn.

2.1 Orthography

It is not hard to see that orthography can play a role in loanword adaptations. For instance, the Afrikaans word *Boer* is adapted as [boɐ] in

French, which is obviously based on the Afrikaans spelling of this word rather than being derived from either its underlying form /bur/ or its surface form [bur]. Of course, very few French speakers would know how to pronounce written Afrikaans words, but even bilingual borrowers have been shown to be influenced by the orthography of the source language, albeit in subtler ways (Vendelin & Peperkamp 2006). For the case at hand, an orthographic account would state that differences in the spelling of word-final [n] in English and French are the source of the observed asymmetry in Japanese. In particular, in English, word-final [n] corresponds most often to the last letter in the written word, whereas in French orthography, it is virtually always followed by the grapheme <e>. Given that in *rōmaji*, the Japanese writing system that uses the Latin alphabet, the moraic nasal consonant is represented by the grapheme <n>, this would explain why loanwords from English are adapted with a final moraic nasal and loanwords from French with an epenthetic vowel.

The obvious problem with this account is that the vowel that is inserted after French word-final [n] is [ɯ], which is represented by <u> in *rōmaji*. Clearly, the vowel [e] would correspond more directly to the grapheme <e> that appears after <n> at the end of French words. It is true that in French a word-final <e> is sometimes realised phonetically as the vowel schwa, for instance in versification or phrase-internally to break up large consonant clusters (as in *centre-gauche* [sɑ̃tʁəgoʃ] 'centre-left'; cf. *centre* [sɑ̃tʁ]); one might thus argue that Japanese bilinguals are aware of the correspondence between <e> and [ə], the latter being most similar to Japanese [ɯ]. This, then, would explain why the epenthetic vowel in French loanwords is [ɯ]. However, [ɯ] is used as the epenthetic vowel not only following word-final [n] in loanwords from French, but also after other coda consonants and within consonant clusters, whether in loanwords from French (*abekku* < *avec* [avɛk] 'with'; *guranpuri* < *grand prix* [gʁɑ̃pʁi]), English (*kurabu* < *club*) or any other language. In the overwhelming majority of these cases, [ɯ] does not correspond to any vowel in the spelling of the source word. It is therefore unlikely that the quality of the epenthetic vowel in the case at hand is related to the relatively weak <e>-to-[ə] correspondence in French. Rather, it seems to have the same motivation (which we will argue below to be phonetic) as in the other cases. We thus reject an orthographic account of the asymmetry between [n]-final loanwords from English and French.

2.2 Phonology

Although in most theories of loanword adaptations the input to the adaptations is constituted by the surface forms of the source language, another view is that adaptations are computed on the basis of a more abstract representation (Paradis & LaCharité 1997, LaCharité & Paradis 2005). According to this view, adaptations are established by highly proficient bilinguals who can access the underlying representation of words in the source language. Differences in the underlying representation of

word-final [n] in English and French, then, could be responsible for the observed asymmetry.

It has indeed been proposed that in French, words ending in a consonant derive from an underlying representation in which this consonant is followed by schwa (Dell 1973); given that no such proposal exists for English, this might explain why loanwords from French but not from English are adapted with a final epenthetic vowel. We reject, however, the hypothesis that adaptations are based on underlying rather than on surface representations, because it makes many wrong predictions. Examples concerning loanwords from English and French are the following. The English word *eroticism*, being derived from *erotic*, contains an underlying /k/; we would thus incorrectly predict the Japanese adaptation, *erochishizumu*, to contain [k] rather than [ʃ]. Likewise, the French adjectives *grand* [gʁɑ̃] 'big' and *petit* [pəti] 'small' contain an underlying final /t/. This so-called liaison consonant surfaces before vowel-initial words only, as in *grand arbre* [gʁɑ̃taʁbʁ] and *petit arbre* [pətitaʁbʁ] 'big/small tree' (Dell 1973). If loanword adaptations were based on underlying representations we would expect the underlying consonant to be present in the Japanese adaptations of these words even before consonant-initial words, contrary to fact. Indeed, we find *guranpuri* from *grand prix*, and *puchiburu* from *petit bourgeois*.[2]

Another possible phonological difference between English and French could lie in the syllabification of word-final consonants. In particular, certain phonological theories allow for these consonants to be syllabified in the onset of an empty-headed syllable, that is, a syllable containing a phonetically empty nucleus (see, for instance, Kaye *et al.* 1990). If word-final consonants are syllabified as onset consonants and hence are followed by a silent vowel in French, this might explain why loanwords from French are adapted with a final vowel. Of course, in order for the asymmetry between loanwords from French and English to be captured, it would have to be shown that in English, word-final consonants are syllabified as proper coda consonants.

It has indeed been proposed that all word-final consonants in French constitute the onsets of empty-headed syllables. Two arguments have been advanced. The first one is phonological: various consonant clusters observed word-finally (such as [pl] at the end of *peuple* [pœpl] 'people') cannot occur in word-internal syllable codas, although they do occur as onset clusters (Dell 1995, Harris 1997, Goad & Brannen 2003). The second argument is phonetic: word-final consonants typically have a vocalic release, regardless of whether they are part of a cluster or not (Goad & Brannen 2003). Concerning English, however, phonological arguments similarly exist in favour of the view that (some) word-final consonants are syllabified as onset consonants. For instance, Goad & Brannen (2003)

[2] One might argue that these expressions have been borrowed via English rather than directly from French. But even if they were first borrowed into English, the problem remains the same: why does the final underlying /t/ from the French sources fail to surface in the English adaptations?

remark that ternary rhymes are excluded in English; word-internally, indeed, syllable rhymes are maximally binary. Therefore, a word-final consonant cannot be part of the preceding syllable if its rhyme already contains a long vowel, diphthong or another consonant (as in *beet*, *five*, *milk*). Goad & Brannen thus propose that such consonants are syllabified as onset consonants. A slightly different stance is taken by Harris (1997). Assuming that within a given language, word-final consonants are uniformly syllabified either as onset or as coda consonants, he proposes that all word-final consonants in English are in the onsets of empty-headed syllables.

Taken together, these proposals make the correct prediction that [n]-final loanwords from French are adapted with an epenthetic vowel, but the incorrect prediction that some or all [n]-final loanwords from English likewise undergo epenthesis. Crucially, we are not aware of any proposal that would make the correct predictions throughout; such a proposal would state that word-final consonants are uniformly syllabified as onset consonants in French and as coda consonants in English. Hence, we reject a phonological analysis based on differences in the syllabification of word-final consonants in English and French.[3]

Yet another phonological explication of the asymmetry in loanword adaptations might be sought in differences in the global phonological systems of English and French. Specifically, it might be argued that proficient bilingual speakers take the phonological grammar of the source language as a whole into account. For the case at hand, the most notable difference in the phonologies of English and French is the status of nasality. In English, nasality is phonemic in consonants, but allophonic in vowels, with oral vowels being nasalised before nasal consonants. In French, by contrast, nasality is phonemic in both consonants and vowels. In particular, the oral vowels /ɛ ɔ ɑ/ have nasal counterparts /ɛ̃ ɔ̃ ɑ̃/. Therefore, words ending in VN sequences can contrast with words ending in Ṽ, as shown by the pair *Cannes* [kan] – *Caen* [kɑ̃] (both cities in France). In Japanese, French nasal vowels are adapted as oral vowels followed by a moraic nasal consonant, as shown in (3).[4]

(3) dessan < dessin [dɛsɛ̃] 'rough sketch'
zubon < jupon [ʒypɔ̃] 'trousers'
roman < roman [ʁomɑ̃] 'spirit of adventure'

[3] Of course, taking only phonetic arguments into account, one might hypothesise that word-final consonants are syllabified as onset consonants in languages such as French, in which these consonants have an important vocalic release, and as coda consonants in the remaining languages, including English. In this case, however, the phonological account of the asymmetry in the adaptation pattern would be nothing other than a more complicated version of the phonetic one, to be discussed in §2.3. That is, phonetic differences in English and French would be responsible for the observed asymmetry, but in an indirect way, via abstract syllable structure.

[4] The glosses are those of the words as they are used in Japanese; they are derived from but do not correspond to the meaning of the source words in French.

If words ending in VN were adapted according to the English pattern, the French contrast between word-final Ṽ and VN would thus be lost. This is shown in (5) with the minimal pair *Caen – Cannes*.

(4) kan < Caen [kɑ̃]
kannu/*kan < Cannes [kan]

Hence this analysis would stipulate that French words ending in VN undergo vowel epenthesis in order to preserve a phonemic contrast present in the source language.[5] The English pattern without epenthesis, by contrast, would be the default adaptation. One would still need to explain, however, why epenthesis occurs even when the final nasal consonant is preceded by a vowel that has no nasal counterpart in French, such as /i/; indeed, the word-final sequence /in/ could very well be adapted as [iN], given that there is no nasal vowel that would yield the same adaptation. Moreover, this analysis has the shortcoming of not being generalisable to many other cases. Loss of foreign phonemic contrasts is indeed massively attested in loanword adaptations. For instance, the oral vowels /u y ø œ/ are all adapted as [u], and both /r/ and /l/ are adapted as [ɾ] in Japanese loanwords from French. Some more examples involving loanwords from French in other languages are: both /r/ and /l/ are adapted as [l] in White Hmong (Golston & Yang 2001), /ʃ/ and /s/ merge to [s] in Fula (Paradis & LaCharité 1997), and word-final /ɑ̃ ɛ̃ ɔ̃/ are adapted as [a e o] in Kinyarwanda, thus yielding the loss of contrast between oral and nasal vowels (Rose 1999). It is hence unlikely that the adaptation of French word-final [n] in Japanese is driven by a requirement to avoid the loss of the French contrast between Ṽ and VN.

2.3 Phonetics

Finally, according to the phonetic account, subtle differences in the realisation of [n] in English and French are held responsible for the observed asymmetry. In particular, word-final [n] typically has a strong release in French but not in English (Tranel 1987). This is the account we defend here. We argue that French and English words ending in [n] are phonetically closest to Japanese forms with and without a final [ɯ] respectively. Hence, whereas the adaptations of final [n] from the two source languages do not involve the same complexity from a phonological point of view, from a phonetic point of view they do: they each represent a phonetically minimal change with respect to the source form to which they apply.

More specifically, we argue that the adaptations of word-final [n] reflect the way in which English and French words are perceived by Japanese listeners. Our reasoning is as follows. Experimental research has shown that during speech perception, systematic distortions apply to all aspects

[5] In this respect, it would be similar to the analysis of the adaptation of English and Japanese [s] into Korean proposed by Ito *et al.* (2006).

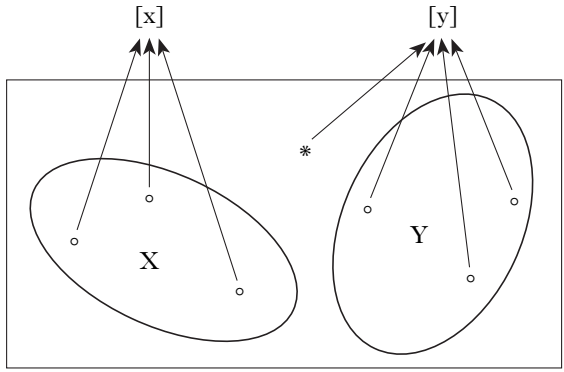

Figure 1
Phonetic decoding of native and non-native sounds. X and Y designate the extensions of the realisation of the segments [x] and [y] belonging to a given language in the acoustic space, represented here as a plane. During perception, native sounds (denoted by ∘) are decoded as the corresponding segment [x] or [y]; the non-native sound (denoted by ∗) is closer to Y than to X and is hence decoded as [y].

of non-native phonological structure, including segments (Goto 1971, Werker & Tees 1984), suprasegments (Kiriloff 1969, Gandour 1983, Lee & Nusbaum 1993, Dupoux *et al*. 1997) and syllable phonotactics (Massaro & Cohen 1983, Hallé *et al*. 1998, Dupoux *et al*. 1999). In particular, non-native sounds and sound structures are perceived as native ones, a phenomenon called PERCEPTUAL ASSIMILATION. This perceptual assimilation of non-native sounds is due to phonetic decoding, a process that transforms continuous universal phonetic representations into discrete language-specific ones (see, for instance, Best 1994). During this process, illustrated in Fig. 1, non-native sounds are assimilated to the closest available phonetic category, with closeness being defined in terms of either acoustic proximity (Kuhl 2000) or proximity in the sense of fine-grained articulatory gestures (Best & Strange 1992).

In order to account for perceptual assimilation of suprasegmental and phonotactic structures, Peperkamp & Dupoux (2003) proposed that the input to the phonetic decoder is constituted by chunks of the acoustic signal having the size of a syllable rather than of a segment. Hence, complete syllables are mapped onto the phonetically closest ones that are well-formed in the listener's native language. Within this framework (see also Peperkamp 2005), loanword adaptations are thus considered to originate in perceptual assimilation (*modulo* a possible effect of orthography) and, therefore, consist of phonetically minimal transformations.[6] Crucially,

[6] An exception is made for loanword adaptations that do not repair illegal phonological structures but represent generalisations to a default pattern in the borrowing language (Shinohara 2000, Kenstowicz & Sohn 2001). For instance, Kenstowicz &

these transformations are not computed by the grammar of the borrowing language, but rather depend upon an acoustic or articulatory distance metric.[7] In this framework, the role of the phonological grammar is restricted to providing the available native forms to which non-native forms are assimilated during phonetic decoding.

Empirical support for the present proposal can be found in a comparison of the loanword literature with the speech-perception literature, showing that psycholinguists have documented several cases of perceptual assimilation that are reflected in loanword adaptations. These perceptual assimilations, moreover, are attested even in highly proficient bilinguals, and they resist specific training. A segmental example is provided by Japanese. This language has only one liquid consonant, and its native speakers find it hard to distinguish the English liquids [ɹ] and [l] in CV stimuli, suggesting that they perceptually assimilate them to their single liquid consonant (Goto 1971). The difficulties experienced by native listeners, which are present also in bilinguals (Flege *et al*. 1996, Aoyama *et al*. 2004) and which resist to a large extent specific training to distinguish [ɹ] and [l] (Lively *et al*. 1993), are reflected in loanwords from English: syllable-initial [ɹ] and [l] are adapted as the same consonant in Japanese (Lovins 1975). A suprasegmental example is provided by French, in which stress predictably falls on the final syllable of the word. French listeners, even those who have a very good mastery of a language with contrastive stress such as Spanish, have severe difficulties perceiving stress contrasts (Dupoux *et al*. 1997, Dupoux *et al*. 2008). This perceptual assimilation effect is reflected in French loanwords, with stress being systematically word-final. Finally, a phonotactic example is again provided by Japanese, in which most consonant clusters are illegal. Both monolingual and bilingual Japanese listeners find it hard to distinguish between non-native clusters and the same clusters broken up by the vowel [u] (Dupoux *et al*. 1999, Nakamura & Dupoux, in preparation), and such clusters undergo vowel epenthesis in loanwords (Lovins 1975).

Sohn (2001) discuss the Korean adaptation of English words like *cinema* with a pitch accent on the penultimate syllable. This adaptation is surprising, since antepenultimate accent would be perfectly acceptable in Korean. A perceptual distortion, then, cannot be the source of the adaptation. Rather, it appears that in words consisting of three light syllables, penultimate accent is the default pattern.

[7] The precise nature of this distance metric remains to be specified. If it is defined in acoustic terms, it should compare the spectral and durational properties as well as the F0 contour of the foreign form to a set of phonotactically legal native forms. A method for computing the distance between spectra, applied by Mielke (2005), consists in converting each spectrum to a matrix composed of a set of Mel-scaled cepstral coefficients computed at short intervals (15 ms), and comparing these matrices by means of Dynamic Time Warping, an algorithm that matches spectrally similar intervals (see, for instance, Rabiner & Juang 1993). In order to make this method appropriate for computing perceptual assimilation of word forms presenting any type of non-native contrast – be it segmental, suprasegmental or phonotactic – the Dynamic Time Warping algorithm should be modified so as to take durational and F0 differences into account as well.

Of course, more experimental work is necessary to establish the correspondence between loanword adaptations and perceptual assimilations. The Japanese treatment of word-final [n] is an interesting test case, since it differs from the previously studied ones in that different source languages give rise to different adaptations of the same phonological structure.

2.4 Summary

To sum up, in order to account for the asymmetry in the Japanese adaptation of word-final [n] in loanwords from English and French, one might propose various hypotheses of an orthographic, phonological or phonetic nature. We have couched the phonetic hypothesis that we favour within a psycholinguistic framework: experimental research with both monolinguals and bilinguals has shown that non-native sound structures are assimilated during perception to the phonetically closest ones that are legal in the native language. We will thus test the hypothesis that loanword adaptations originate in perceptual assimilation. That is, rather than directly testing the phonetic nature of the loanword adaptations by carrying out extensive measurements on English, French and Japanese stimuli, we will examine if the Japanese loanword data mirror the way in which Japanese listeners perceive word final [n] in English and French. If this is the case, then we will conclude, following models of perceptual assimilation proposed by Best (1994), Kuhl (2000) and Best & Strange (1992), that the adaptations of English and French word-final [n] in Japanese are based on phonetic minimality.

In the first experiment, we use an offline identification task with English and French non-word stimuli; by using monolingual Japanese participants, we make sure that neither the orthography nor the phonology of English and French can play a role. In the second experiment, we use an online discrimination task with French non-word stimuli and bilingual Japanese-French participants. Although these bilinguals know the French spelling rules and might have some detailed knowledge of its phonology, the choice of the task ensures that neither one of these sources of knowledge can be put to use during the experiment. Indeed, phonetic decoding, the processing level that is responsible for perceptual assimilations and that is tapped by online discrimination tasks, precedes both orthographic and phonological processing levels (see, for instance, Dehaene-Lambertz *et al*. 2000).

3 Experiment 1

Our experimental investigation tests the hypothesis that English forms ending in [n] are perceptually closest to Japanese forms ending in a moraic nasal consonant, whereas French forms ending in [n] are perceptually closest to Japanese forms ending in an onset [n] followed by a

vowel.[8] In particular, we administer a forced choice identification task with non-word stimuli produced by American English and French speakers. These stimuli are presented orally to Japanese listeners who have very little or no knowledge of French and limited knowledge of English, and who are, moreover, not informed about the English and French nature of the stimuli. This design thus guarantees that any differences observed in the responses to the English and French stimuli can only be due to differences in the phonetic properties of the stimuli, and not to differences in the orthographic or phonological properties of English and French.

3.1 Stimuli

We created 25 trisyllabic items of the form CVCV'CVN, containing the following vowels[9] and consonants: [i u ɛ ɔ ɑ p t k b d g m n s j] (see Appendix). All items respected the phonotactic structure of English and French, and were non-words in both these languages. Twenty items were embedded in both an English and a French carrier sentence, in which they appeared as invented proper names. In the English sentences, these names were composed of a disyllabic first name and a monosyllabic last name, such that main stress was on the last syllable (e.g. *Mahby Yoon* for the item [mɑbi'jun]). French has fixed word-final stress, and in the French sentences, the proper names were trisyllabic first names (e.g. *Mabiyoune*). The English and French carrier sentences were matched in the number of syllables, the position of the non-word item and the initial consonant of the word to its right. The English sentences were read by four native speakers of American English (two men and two women) and the French ones by four native speakers of French (two men and two women).[10] The remaining five items were embedded in a Russian carrier sentence and read by a female native speaker of Russian; they were to be used for a short training. All sentences were recorded on a DAT-recorder, digitised at 16000 Hz, and stored on a computer disk. The CVCV'CVN non-words were then sliced out on the basis of inspection of the waveforms and spectrograms, and stored as individual sound files. We thus obtained eight stimuli – four English and four French ones – for each of the 20 test items, and a single stimulus for each of the five training items.

Figure 2 shows spectrograms of the item /dɛbi'pin/, spoken by a male and a female English speaker and a male and a female French speaker. The segmentation of the stimuli was determined by means of both auditory

[8] A pilot experiment with similar results was presented in Vendelin & Peperkamp (2004). This pilot used the same task but different stimuli.

[9] We chose vowels that are relatively similar in French and English. Some of these do not exist in Japanese, but this is not a problem in a forced choice task that does not focus on vowel perception.

[10] In French, [ɑ] is very infrequent, and the French speakers produced this vowel most often as [a]. Note also that none of the French speakers had a Southern French accent, where final consonants are often followed by a schwa.

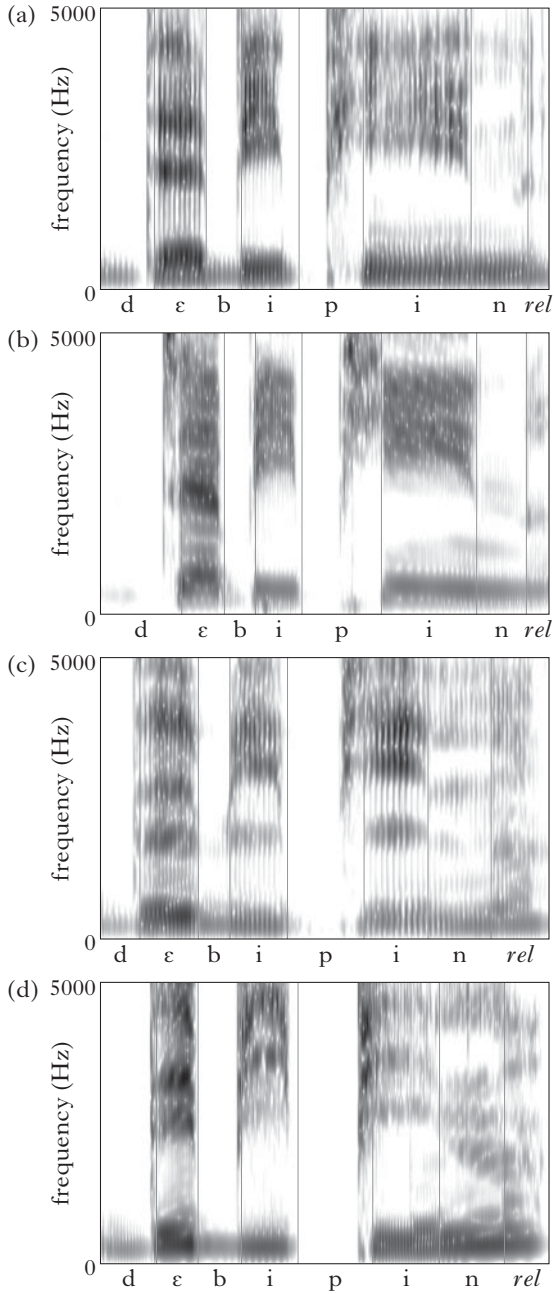

Figure 2

Spectrograms of the item /dɛbi'pin/, spoken by (a) a male English speaker, (b) a female English speaker, (c) a male French speaker and (d) a female French speaker. They are normalised for duration and have a dynamic range of 40 dB (*rel* = release).

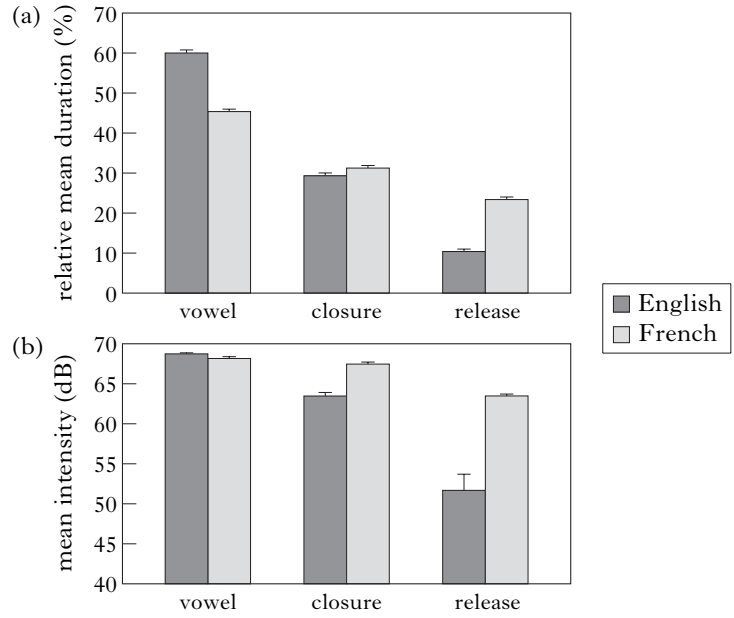

Figure 3

Relative durations (a) and mean intensities (b) of the vowel, the nasal closure, and the release in the final rhyme of the English and French stimuli.

and visual inspection. In particular, the presence of a vocalic release is revealed by a transition from a concentration of spectral energy at the lower frequencies – characteristic of nasal consonants – to a more homogenous redistribution of the spectral energy.

The mean durations of the American English and the French stimuli were 841 ms and 715 ms respectively, representing a significant difference ($t(158) = 13{\cdot}2$, $p < 0{\cdot}0001$). The duration of the final VN rhymes was also significantly different, with the English ones lasting longer than the French ones (English: 370 ms; French: 310 ms; $t(158) = 10{\cdot}2$, $p < 0{\cdot}0001$). In the English stimuli, the nasal consonant was released in 90% of the cases and in the French ones in 100% of the cases, again a significant difference ($t(158) = 3{\cdot}0$, $p < 0{\cdot}005$). In the English stimuli, the releases produced by the female speakers had vocalic formants, whereas those produced by the male speakers were better characterised as aspiration. In the French stimuli, all releases had vocalic formants. Figures 3a and b show the relative durations and mean intensities respectively of the vowel, the nasal closure and the release in the final rhyme of the English and French stimuli. There were significant differences in the relative durations of the vowel ($t(158) = 15{\cdot}6$, $p < 0{\cdot}0001$), the nasal closure ($t(158) = 2{\cdot}0$, $p < 0{\cdot}05$) and the release ($t(158) = 14{\cdot}6$, $p < 0{\cdot}0001$). Furthermore, there were significant differences in the mean intensities of the vowel

($t(158) = 2{\cdot}2$, $p < 0{\cdot}03$), the nasal closure ($t(158) = 7{\cdot}8$, $p < 0{\cdot}0001$) and the release ($t(158) = 5{\cdot}8$, $p < 0{\cdot}0001$).

Finally, among the French speakers, the releases of the male speakers had a longer relative duration than those of the female speakers (male: 26·3 %; female: 20·5 %; $t(78) = 5{\cdot}2$, $p < 0{\cdot}0001$), but they did not differ in their mean intensities ($t < 1$). Among the American English speakers, there were no differences in the relative duration and intensity of the releases produced by the male *vs.* the female speakers.

3.2 Procedure

For the main part of the experiment, we created 20 blocks, one per item, each consisting of four American English and four French tokens. Tokens were randomly shuffled within the blocks and separated by five-second silences. Participants completed a forced choice identification task: after listening to a token, they had to indicate on an answer sheet the candidate non-word that in their opinion was closest to the token they had heard. At the end of each block, participants had to press a button to start listening to the next block. The main part of the experiment was preceded by a short training phase, in which participants performed the same task on the five training stimuli.

The answer sheet contained six options for each token. These options – always non-words – were identical as far as the first two syllables were concerned, but differed in what followed. Given that both vowel length and consonant length are phonemic in Japanese, the set of answer options resulted from combining three binary parameters: simple *vs.* geminate vowel in the third syllable, simple *vs.* geminate nasal in the third syllable, and absence *vs.* presence of a final vowel. The second parameter is dependent on the third one, since a geminate consonant can only occur if followed by a vowel. For each item, there were thus two options without and four with a final vowel. An example is shown in (5). On the answer sheet, all options were written not in *rōmaji* as in (5), but in *katakana*, the Japanese syllabary used for loanwords.

(5) *Forced choice options for the item* [mɑbiˈjun]
mabiyun – mabiyūn – mabiyunu – mabiyūnu – mabiyunnu – mabiyūnnu

Note that the final vowel was always <u>, corresponding to the vowel [ɯ]. This is indeed the default epenthetic vowel in both integrated loanwords and online adaptations in Japanese; moreover, it is the vowel that Japanese speakers perceive within non-native consonant clusters (Dupoux *et al.* 1999).

From the viewpoint of phonological well-formedness, two of the options are marginal, since they contain a trimoraic syllable; this holds for the second one, containing a long vowel followed by a tautosyllabic nasal, and the last one, containing a long vowel followed by a geminate nasal. With a handful of exceptions (e.g. *pātto* 'swiftly'), the Japanese core

lexicon does not allow for trimoraic syllables. In onomatopoeias and loanwords, however, words ending in a sequence V:N freely occur (e.g. *kīn* 'shrill metallic sound', *supūn* 'spoon') (Kubozono 1999).

At the end of the experiment, participants were told that they had been listening to speakers of two different languages; they were then asked to guess which languages they had heard.

The experiment lasted about 25 minutes.

3.3 Participants

Twelve native speakers of Japanese (nine men and three women), with no known hearing problems, aged between 20 and 32 (mean 24·4), were tested in Tokyo. Eleven participants had learned English in school, starting at age 12 or 13. Self-judgements on a ten-point scale of their competences in English were on average 5·6 for comprehension, 4·5 for production and 4 for pronunciation. The remaining participant had no knowledge of English at all. Three participants had learned French, one starting at age 8, one at age 19 and one at age 20; all three judged their own competence in both comprehension and production in this language as well as their pronunciation as very bad, between 1 and 4 on a ten-point scale (average 1·7 for comprehension, 1·7 for production and 1·8 for pronunciation).

Five participants correctly identified English as one of the languages that were used in the experiment, and nine participants correctly identified French (three participants identified both English and French). Other languages that were mentioned were German (two participants), Italian (two), Spanish (two), Hindi (one), Russian (one), Turkish (one) and Urdu (one).

3.4 Results and discussion

The mean percentages of responses with a final vowel, separated by speaker, are shown in Table I.

	English					French				
speaker	A	B	C	D	*mean*	E	F	G	H	*mean*
% responses with a final vowel	45·8 (6·2)	52·9 (6·0)	19·6 (2·2)	30·4 (5·3)	37·2 (4·3)	72·5 (6·1)	64·2 (6·3)	97·5 (1·2)	76·7 (7·3)	77·7 (4·5)

Table I

Mean percentages of responses with a final vowel and Standard Errors (between parentheses) for English and French stimuli, separated by speaker. Speakers A, B, E and F are female; the others are male.

The mean percentages of responses that included a final vowel were submitted to a one-way ANOVA with the within-participant factor

Language (English *vs.* French). The effect of Language was highly significant ($F(1, 11) = 170{\cdot}2$, $p < 0{\cdot}0001$). Next, the same data were submitted to two one-way ANOVAs, restricted to the English and the French stimuli respectively, with the within-participant factor Sex (male *vs.* female). The effect of Sex was significant in both ANOVAs (English: $F(1, 11) = 74{\cdot}8$, $p < 0{\cdot}0001$; French: $F(1, 11) = 24{\cdot}2$, $p < 0{\cdot}0001$). Finally, a series of regression analyses was carried out to examine which acoustic characteristic(s) of the stimuli best predicted the presence of a final vowel in the responses. The following variables were used as predictors: relative duration of vowel (i.e. the vowel preceding the nasal consonant), relative duration of nasal, relative duration of release, intensity of vowel, intensity of nasal, intensity of release, relative duration of vowel multiplied by its intensity, relative duration of nasal multiplied by its intensity and relative duration of release multiplied by its intensity.[11] All relative durations were calculated within the final rhyme. It was found that the presence of a final vowel in the responses was best predicted by a single variable, the relative duration of the release multiplied by its intensity, in other words the spectral energy contained within the release of the nasal consonant (see Fig. 4).[12] This variable alone accounted for 61% of the variation ($F(1, 158) = 242{\cdot}1$, $p < 0{\cdot}0001$).

These results show that for native speakers of Japanese, [n]-final items are perceptually closest to a nasal consonant followed by a vowel significantly more often when produced by French speakers than when produced by American English speakers. In particular, in stimuli produced by French speakers, the participants in this experiment identified the final [n] as a Japanese consonant–vowel sequence in three-quarters of the cases, while in stimuli produced by American English speakers, they did so in about only one-third of the cases. Among the English stimuli, most responses with a final vowel were given to those produced by female speakers, that is, to the stimuli whose final release more often had vocalic formants. By contrast, among the French stimuli, most responses with a

[11] Another factor of interest is the amount of nasalisation on the vowel. English and French differ in this variable in that, as mentioned in §2.2, vowels are nasalised before nasal consonants in English, but not in French. Given that in Japanese, the moraic nasal consonant can be realised as a nasal copy of the preceding vowel, we expect that the presence of vowel nasalisation induces Japanese listeners to choose a response without a final vowel. However, in the absence of a measure of nasal airflow in our stimuli, it was impossible to enter this variable as a predictor.

[12] A reviewer wonders why we do not obtain a sigmoid curve, characteristic of categorical perception. Such a curve basically consists of three parts: a low horizontal line (floor effect), a tilted line, as found in our regression, and a high horizontal line (ceiling effect). Concerning the rightmost part of the graph, it is likely that we would have found the ceiling effect with stimuli with an even higher amount of spectral energy. Concerning the leftmost part of the graph, given that there were stimuli having no release at all, the absence of the floor effect is unexpected at first sight. Note, however, that before pause, the Japanese moraic nasal is produced most often as a uvular nasal or a vowel, not as a dental or alveolar [n]. In other words, even unreleased tokens of word-final [n] do not match the Japanese word-final N very well and rather tend to be identified as the sequence [nɯ].

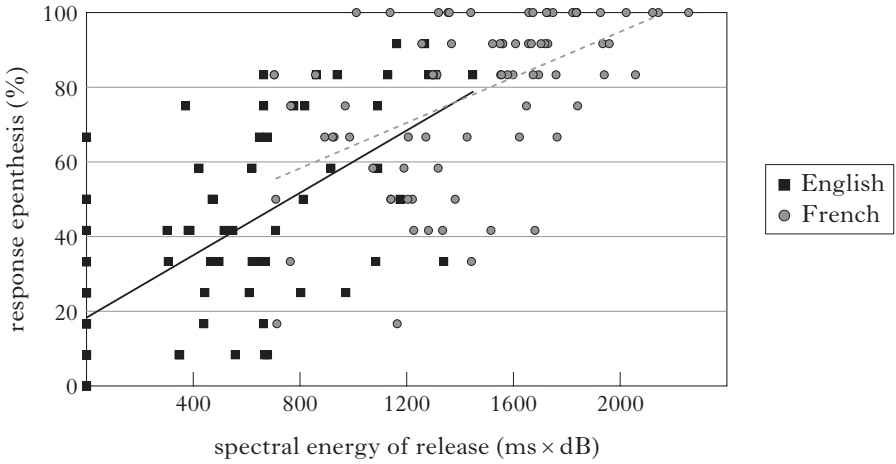

Figure 4
Mean percentages of responses including an epenthetic vowel to tokens spoken by English and French speakers as a function of the spectral energy of the release of the nasal consonant. The straight and dotted linear regression lines concern the English and French tokens respectively.

final vowel were given to those produced by male speakers, that is, to the stimuli whose final release on average had a longer relative duration. Overall, the presence of the final vowel in the responses correlated positively with the amount of spectral energy contained in the release of the nasal consonant. These identification judgements are uncontaminated by either orthographic or phonological influences. This is because, firstly, the presentation of the stimuli was auditory and the participants were not even informed about the English and French nature of the stimuli, and, secondly, the participants had little or no knowledge of French and only limited knowledge of English.

Hence Japanese speakers are sensitive to phonetic differences between English and French realisations of word-final [n], and the different response patterns found with the identification task mirror those found in loanword adaptations. However, these results do not show that Japanese speakers necessarily perceive English word-final [n] as their native moraic nasal consonant and French word-final [n] as a sequence of nasal consonant + vowel respectively. This is because the options to choose from were limited to a number of forms that are legal in Japanese. For instance, it might be the case that the participants perceived a vocalic release rather than a full-fledged vowel in the French stimuli, a release that they felt was most similar to the Japanese vowel [ɯ]. In order to show that the presence of an epenthetic vowel in loanwords from French originates from perceptual assimilation (that is, mirrors a transformation that takes place during phonetic decoding), it is necessary to go one step further. Specifically, we should show that Japanese speakers have difficulties

perceiving French word-final [n] as anything other than consisting of a nasal consonant followed by their native vowel [ɯ]. This is the aim of the next experiment, in which we use a discrimination rather than an identification task.

4 Experiment 2

Perceptual assimilation is assessed in online discrimination experiments, using for instance an ABX paradigm. In this paradigm, A and B are always different and X is identical to either A or B; the task is to indicate which one of A and B is identical to X. The harder it is to discriminate between A and B, the lower the participants' performance. Ideally, we would use this task to assess the discrimination between French stimuli ending in [n] and corresponding Japanese stimuli with an additional final [ɯ]. However, given that French and Japanese differ on many dimensions, it would be very easy for Japanese participants to perform the task without paying attention to the final part of the stimuli. Indeed, among A and B, one would sound Japanese and the other one distinctly non-Japanese. A fruitful response strategy would thus consist in choosing the one that sounds Japanese or the one that does not, according to whether X itself sounds Japanese or not; a comparison restricted to the final part of the stimuli would simply be unnecessary.[13]

In this experiment, we circumvent this problem by using French stimuli only. That is, we assess the discrimination of pairs of non-words, produced by French speakers, which differ only in that they end either in [n] or in [nu] (e.g. [mabijun] – [mabijunu]). We reason that if Japanese speakers, as opposed to French ones, have difficulties perceiving the difference, this is evidence that both non-words are perceptually assimilated to a single native form, ending in [ɯ].[14] As a control condition, we also assess the discrimination of pairs of non-words such as [mabijun] – [mabijuni], where the difference lies in the presence *vs*. absence of a final [i]. For this contrast, Japanese speakers should have no more difficulty than French speakers, regardless of whether they perceive a final [ɯ] in [mabijun] or not.

4.1 Stimuli

Using the same items as in Experiment 1, we created 25 triplets of the form [CVCVCVn] – [CVCVCVnu] – [CVCVCVni], five to be used for a

[13] The same problem holds for a different discrimination task, AX, where participants have to indicate if X is identical to A or not. That is, a response strategy that yields only correct responses would consist in replying 'yes' if both A and X sound either Japanese or non-Japanese and 'no' otherwise.

[14] Note that in loanwords from French, the vowel [u] is adapted as [ɯ] (examples are *konkūru*, adapted from *concours* [kɔ̃kuʁ] 'contest', and *ōto-kuchūru*, adapted from *haute couture* [otkutyʁ] 'high fashion').

training phase and the remaining 20 for the test phase. Two native French speakers with some phonetic training, a man and a woman, recorded the triplets. They were instructed to put stress on the vowel preceding [n] (yielding non-French penultimate stress in the case of the vowel-final tokens). All stimuli were recorded on a DAT recorder, digitised at 16000 Hz and stored on a computer disk.

For each item, we compared the [n]-final token produced by the male speaker to the one produced by the female speaker, and, where necessary, digitally edited the duration of the prenasal vowel, the consonant [n] and the final vowels [i] and [u], such that Japanese listeners would not perceive any differences in vowel or consonant length between the two tokens produced by the different speakers. Next, for each speaker and for each triplet, we extracted the initial CVCVCV portion from the [CVCVCVn] token and the final [nu] and [ni] portions from the vowel-final tokens. We then created new vowel-final tokens by adding the [nu] and [ni] endings to the CVCVCV base. In the resulting triplets ([CVCVCVn] – [CVCVCVnu] – [CVCVCVni]), the differences between the three tokens thus concerned the realisation of the [n], [nu] and [ni] endings only, their initial CVCVCV portions being strictly identical. Finally, for each item we compared the duration of the final vowel or vocalic release across the six tokens (three male and three female), and, where necessary, reduced or enhanced the duration of [i] or [u] in one or more tokens, such that Japanese listeners would not perceive any length differences. For each triplet, the resulting six stimuli were thus homogeneous in terms of both the F0 contour and the segmental durations, making it impossible for the Japanese participants to discriminate the vowel-final stimuli from the [n]-final ones on the basis of perceived differences in either pitch accent or vowel or consonant length. (For French participants, this was hardly a risk to begin with, neither stress nor consonant and/or vowel length being contrastive in French.)

Finally, a native speaker of French listened to the stimuli and verified that they all ended in the intended segment (i.e. [n], [i] or [u]).

4.2 Procedure

Participants were told that they would listen to sequences of three words in a foreign language, that in each sequence the third word was the same as either the first or the second one, and that their task would be to indicate which of the first two words was the same as the third one. Each experimental trial consisted of the presentation of three stimuli belonging to the same triplet (A, B and X), with an interstimulus interval of 500 ms. A and B were always produced by the female speaker and X by the male speaker. Participants had 3500 ms to press a button on their left or right to indicate whether X was the same as A or B respectively. The trial ended immediately after the response was given or after the 3500 ms had elapsed, whichever came first. The next trial started 1000 ms later.

The experiment started with a training phase of 20 trials, during which participants received feedback as to whether their responses were correct. In the case of an incorrect response or no response within 3500 ms the trial was repeated until the correct response was given. The test phase consisted of 160 trials, 8 per item, divided over two blocks. In each block, half of the trials concerned the experimental contrast [n] *vs.* [nu], the other half the control contrast [n] *vs.* [ni]; the quality of X (ending in [n] or in a vowel) as well as the correct response (A or B) were counterbalanced. Participants could take a short pause after the first block.

The experiment lasted about 20 minutes.

4.3 Participants

Sixteen native speakers of French, four men and 12 women, and 16 native speakers of Japanese, three men and 13 women, were tested individually in Paris. The French participants were aged between 18 and 28 (mean 22). The Japanese participants were aged between 21 and 37 (mean 29); they had lived in Paris between one and seven years (mean three). They had started to learn French between 12 and 26 years of age (mean 19·6), mainly at university and/or in a language school. Self-judgments on a ten-point scale of their own competences in French were on average 6·3 for comprehension, 5·9 for production and 5·9 for pronunciation. None of the participants reported a known hearing problem.

4.4 Results and discussion

One French participant had a correct response rate of 26 %, i.e. well below the level of chance, which is at 50 %. This suggests a confusion about the association of the response buttons; the data from this participant were therefore discarded.

Figure 5 shows the mean correct response rates and the mean reaction times for the trials with a correct response, as a function of native language and type of contrast.

Two ANOVAs were run, one on the correct response rates and one on the reaction times, with the between-participants factor Language (Japanese *vs.* French) and the within-participant factor Contrast ([n] – [nu] *vs.* [n] – [ni]). The ANOVA on the correct response rates revealed significant main effects of both Language ($F(1, 29) = 38{\cdot}6$, $p < 0{\cdot}0001$) and Contrast ($F(1, 29) = 63{\cdot}8$, $p < 0{\cdot}0001$), as well as a significant interaction between these two factors ($F(1, 29) = 38{\cdot}4$, $p < 0{\cdot}0001$). This interaction was due to the fact that the size of the effect of Contrast was much bigger for the Japanese than for the French participants (Japanese: 72·0 % *vs.* 94·8 %, $F(1, 15) = 55{\cdot}9$, $p < 0{\cdot}0001$; French: 95·7 % *vs.* 98·3 %, $F(1, 14) = 8{\cdot}9$, $p < 0{\cdot}02$). The ANOVA on the reaction times for the trials with a correct response revealed a marginally significant effect of Language ($F(1, 29) = 3{\cdot}3$, $p < 0{\cdot}078$), a significant effect of Contrast ($F(1, 29) = 36{\cdot}8$, $p < 0{\cdot}0001$) and a significant interaction between these two factors ($F(1, 29) = 16{\cdot}9$,

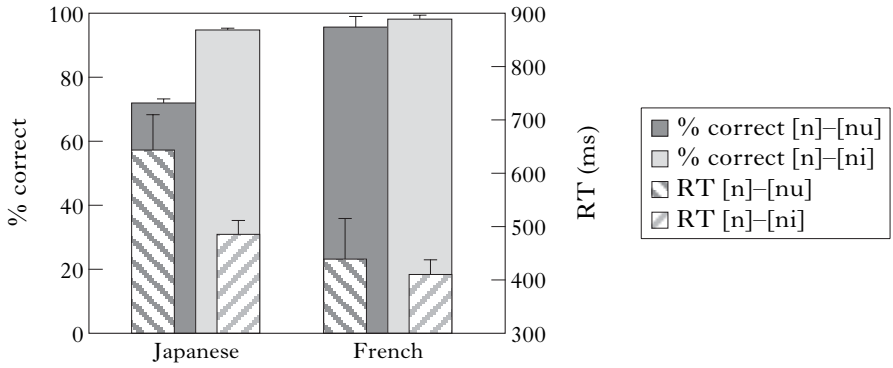

Figure 5

Mean correct response rates (large bars) and mean reaction times (small, striped bars) for the trials with a correct response, as a function of native language and type of contrast. The error bars indicate standard errors.

$p < 0{\cdot}0001$). This interaction was due to the fact that the effect of Contrast was significant for the Japanese (643 ms *vs.* 485 ms, $F(1, 15) = 35{\cdot}6$, $p < 0{\cdot}0001$), but not for the French participants (439 ms *vs.* 410 ms, $F(1, 14) = 3{\cdot}2$, $p = 0{\cdot}1$).

These results show that Japanese listeners who are proficient in French are both more error prone and slower to discriminate French [n]-final stimuli from corresponding [nu]-final ones than from corresponding [ni]-final ones. French listeners also make more errors on the [n] – [nu] contrast (suggesting that the schwa-like vocalic release of [n] is acoustically closer to [u] than to [i]), but the difference is very much reduced compared to that obtained with Japanese listeners; moreover, French listeners respond equally fast to the two contrasts. This, then, is evidence that Japanese listeners, even bilingual ones who live in France, have difficulties perceiving the French contrast between the vocalic release of final [n] and the vowel [u], which are hence perceptually assimilated to the same Japanese vowel, i.e. [ɯ].

Note that the performance of the Japanese participants in this experiment is still better than chance. This is not surprising, given the presence of acoustic differences between the [n]-final stimuli on the one hand and the [nu]-final ones on the other hand. Better than chance performance is indeed widespread in experiments concerning even the most difficult non-native contrasts, such as the English [ɹ] – [l] contrast for Japanese listeners (Goto 1971) or the Spanish stress contrast for French listeners (Dupoux *et al.* 1997). In particular, several studies have shown that the more the phonetic variability of the stimuli is reduced, the easier it is to perceive non-native contrasts (see, for instance, Werker & Tees 1984, Dupoux *et al.* 1997, Dupoux *et al.* 2001). In all cases of perceptual assimilation, the crucial finding is that listeners fail to phonologically encode acoustic differences that do not represent a phonemic contrast in their native

language, as shown by the fact that their performance is significantly worse than that of a control group.[15] Following Werker & Tees (1984), we argue that for the purposes of loanword adaptation the residual acoustic capacity to perceive non-native contrasts is of little use. This is because candidate loanwords are typically embedded in sentences and pronounced by a large variety of speakers, yielding a much higher amount of phonetic variability than in even the most difficult perception experiment. Moreover, the noisy environment in which natural language processing occurs adds yet another dimension to the difficulty of accurate perception of non-native contrasts.

5 General discussion

Taken together, the results of the present experiments show that Japanese listeners perceptually assimilate French but not English word-final [n] to a consonant-vowel sequence. This is strong evidence that the asymmetry found in the adaptations of [n]-final loanwords from French and English originates in phonetic decoding during speech perception, a process that has been argued to be based on phonetic minimality, expressed in terms of either acoustic parameters (Kuhl 2000) or fine-grained articulatory gestures (Best & Strange 1992). Ultimately, phonetic measurements comparing Japanese productions to English and French ones are necessary to confirm that the perceptual assimilation of English and French word-final [n] consist of phonetically minimal transformations.[16]

In this section, we discuss how the present research relates to previous work on the role of perception in loanword adaptations. Specifically, we

[15] The question of acoustic *vs*. phonological processing has been studied in more detail for a case that is very similar to the one presently investigated. Japanese listeners, contrary to French ones, have difficulties perceiving contrasts such as that between [ebzo] and [ebuzo], where the only difference lies in the absence *vs*. presence of the vowel [u]; by contrast, French listeners were shown to have difficulties perceiving vowel-length contrasts, e.g. [ebuzo] – [ebu:zo], which pose no problem to Japanese listeners (Dupoux *et al*. 1999). Using functional Magnetic Resonance Imaging (fMRI), Jacquemot *et al*. (2003) next showed that the contrast that is phonologically irrelevant and hence difficult to perceive (i.e. [ebuzo] – [ebu:zo] for French listeners and [ebzo] – [ebuzo] for Japanese listeners) fails to activate regions in the left hemisphere that are activated by the contrast that is phonemic and hence easy to perceive (i.e. [ebuzo] – [ebu:zo] for Japanese listeners and [ebzo] – [ebuzo] for French listeners). Interestingly, participants performed a very easy task in this fMRI study: an AAX discrimination paradigm was used in which the two occurrences of A were physically identical and X was either the same token once again or a token of the contrastive category. Consequently, participants could perform the task at an acoustic processing level and indeed did quite well on the non-native contrast. Yet, the brain-imaging data showed a large processing advantage for the native contrast.

[16] Not surprisingly, it turns out that a similar asymmetry between loanwords from English and French can be found in other languages. In particular, Kang (1996) reports that in Korean, a word-final [m] – a legal coda consonant – is adapted with a final epenthetic vowel in loanwords from French but not in loanwords from English. As in the present paper, Kang invokes the strong vocalic release of word-final consonants in French to account for this asymmetry.

first review experimental and phonetic data, and then turn to a comparison of our psycholinguistic modelling of the data with alternatives couched within phonological theory.

5.1 Previous work

Following the seminal article of Silverman (1992), the possible role of perception in loanword adaptations has received a great deal of attention (Yip 1993, 2002, Takagi & Mann 1994, Rose 1999, Gbéto 2000, Kim & Curtis 2002, Kang 2003, Broselow 2004a, Kenstowicz 2004, Iverson & Lee 2006, Smith 2006, Davidson 2007; see also several of the contributions to Kenstowicz & Uffmann 2006 for arguments both for and against the role of perception). It seems fair to say that nowadays there is some consensus that perception plays a role in loanword adaptations. There is less agreement, however, as to the extent to which perception is responsible for adaptation patterns. Of course, this is ultimately an empirical question. Whereas arguments in favour of perception are often based on impressionistic data, experimental data have been presented as well, in particular by Takagi & Mann (1994) and Kim & Curtis (2002). The present results are in line with both of these previous reports concerning the perceptual origin and phonetic character of loanword adaptations.

Takagi & Mann (1994) examined the systematic differences in Japanese adaptations of English words containing a vowel + stop sequence as a function of whether the vowel is tense or lax. The basic pattern is that word-finally, the VC sequence is adapted as a short vowel followed by a geminate consonant (and an epenthetic vowel) if the English vowel is lax (6a) and as a long vowel followed by a simple consonant (and an epenthetic vowel) if the English vowel is tense (6b).[17]

(6) a. hitto < hit b. bītō < beat

Following up on work by Lovins (1975), Takagi & Mann conducted a perception experiment that showed that this asymmetry is reflected by the way in which Japanese listeners, both monolingual and bilingual, identify English vowel + stop sequences. That is, English non-word stimuli ending in such a sequence were judged to be closest to a Japanese sequence consisting of a short vowel plus a geminate consonant if the English vowel was lax and of a long vowel plus a simple consonant if it was tense.

Similarly, Kim & Curtis (2002) examined the Korean adaptation of English /s/ as lenis /s/ if it is followed by a consonant (7a) and as fortis /s'/ in other contexts, i.e. before a vowel (7b) or word-finally (7c).

(7) a. [**s**ɨlɤmp^{h}ɨ] < slump
    b. [**s'**olo] < solo
    c. [ʧen**s'**ɨ] < chance

[17] Note that in these and other loanwords ending in [t], the epenthetic vowel is [o] rather than [ɯ]. The sequence [tɯ] is not allowed in Japanese.

Although duration is not the main cue to lenis/fortis contrasts in Korean, the frication duration of lenis /s/ is shorter than that of fortis /s'/. Conducting both a production and a perception experiment, Kim & Curtis tested the hypothesis that the adaptation pattern in (7) mirrors Korean speakers' perception of subtle differences in the duration of /s/ in English. In the production experiment, they showed that English /s/ is indeed shorter before a consonant than in other contexts. In the perception experiment, they digitally edited an English syllable [sa], creating several tokens that varied in the duration of [s], and asked Korean listeners to identify the consonant as either lenis or fortis. It was found that there is a linear relationship between the duration of [s] and the percentage of lenis responses; that is, the shorter its duration, the more English [s] was judged to be similar to lenis /s/. Although this experiment does not directly test the perception of English [s] in different contexts, it shows that Korean listeners are sensitive to the durational differences that English [s] displays.

The present article, as well as Takagi & Mann (1994) and Kim & Curtis (2002), involves direct comparisons of loanword adaptations to the perception of non-native sound structures. A study that is slightly different in character but that should not be left unmentioned is that of Kang (2003). Without arguing that loanword adaptations actually originate in speech perception, Kang hypothesises that they are driven by a requirement to achieve maximal perceptual similarity between the source forms and their adaptations. Equating perceptual similarity with phonetic proximity, she uses phonetic measurements to provide evidence in favour of this hypothesis. Her study deals with English loanwords in Korean, in particular those that end in a stop. These stops sometimes undergo what seems to be an unnecessary adaptation. For instance, voiceless stops are sometimes adapted as an aspirated stop followed by an epenthetic vowel, despite the fact that in Korean voiceless stops are perfectly legal word-finally. Kang shows that the occurrence of this adaptation depends upon several factors, one of which is the nature of the preceding vowel: epenthesis is indeed likely to occur if this vowel is tense, as in the example in (8a), but not if it is lax, as in (8b).

(8) a. [wi**k**$^{\mathbf{h}}$**i**] < week b. [k$^{\mathrm{h}}$wi**k**] < quick

From a phonological point of view, this pattern is unexpected. From a phonetic point of view, however, it makes sense. That is, in Korean, final stops are obligatorily unreleased, whereas in English, they can be realised either with or without a release. In the former case, the phonetically closest match is with a Korean form that has a final vowel. Crucially, when preceded by a tense vowel, English stops appear to be more often released, whereas when preceded by a lax vowel, they are more often unreleased, a pattern that matches the distribution of epenthesis in loanword adaptations.

Interestingly, the Korean cases studied by Kim & Curtis (2002) and Kang (2003) show a certain similarity to the Japanese one considered in

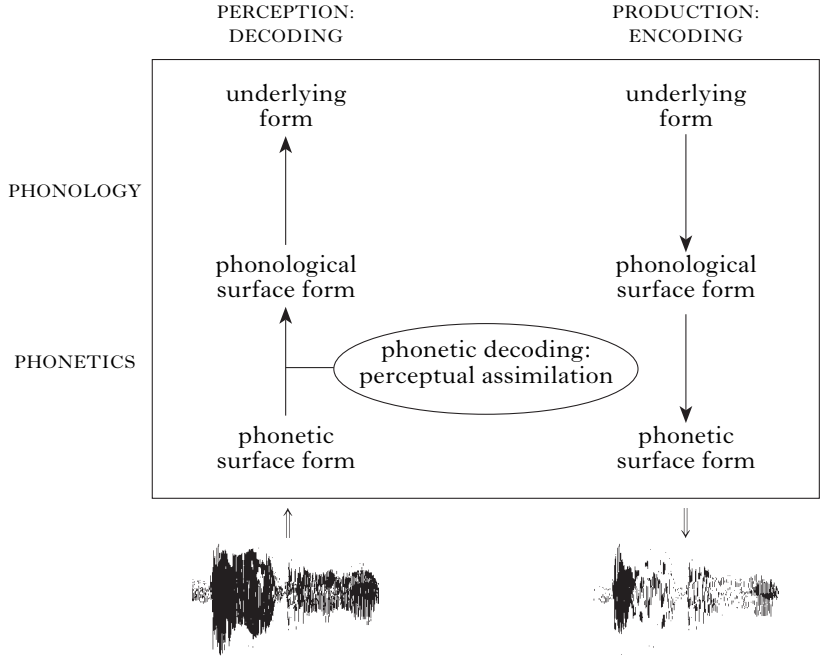

Figure 6
A speech-sound processing model. The processes indicated by single arrows are language specific; those marked by double arrows are universal.

the present article: a single segment gives rise to two different adaptation patterns within the same borrowing language, with the choice of the adaptation depending upon the phonetic realisation of that segment in the input. In the Korean cases this realisation differs as a function of the phonological context, whereas in the Japanese case it differs as a function of the source language.

5.2 Psycholinguistic *vs*. phonological modelling

More experimental work is necessary to test the perceptual origin of loanword adaptations. In this section, we turn away from this empirical question and examine how the role of perception might be modelled. In §2.3, we proposed that loanword adaptations reflect the process of perceptual assimilation, which takes place during phonetic decoding. Figure 6 shows the place of phonetic decoding within an integrated speech-sound processing model. This model, based on various proposals from the literature on speech perception and production, makes two orthogonal distinctions: one between perception and production, the other one between phonology and phonetics.[18] During perception, an

[18] Outside the realm of loanword adaptations, Boersma (1999) proposes a grammar model that likewise distinguishes between perception and production and recognises

acoustic signal is first of all mapped onto a phonetic surface form, based on a continuous, universal format of formant transitions. This universal phonetic representation is then mapped onto a language-specific discrete phonological surface form by the process called phonetic decoding (Best 1994); recall that this is the process responsible for perceptual assimilation of non-native sounds and sound structures to native ones. Finally, the surface phonological form is mapped onto an underlying form, a process called phonological decoding (Lahiri & Marslen-Wilson 1991). During production, on the other hand, the phonological encoder maps underlying forms onto surface phonological forms, which are mapped onto surface phonetic forms by the phonetic encoder (Levelt 1989, Levelt *et al*. 1999).

The underlying form in perception is homologous to the underlying form in production; likewise for the two phonological surface forms. The phonological decoder and encoder thus perform inverse functions. The phonetic surface forms can diverge quite a bit, especially in the case of the processing of non-native words, where the phonetic surface form in perception is a low-level representation in terms of formant transitions of an acoustic signal produced in a foreign language and the one in production contains the articulatory motor program for its nativised form. Our proposal that loanword adaptations reflect the phenomenon of perceptual assimilation formalises an intuition expressed in Silverman (1992) and Yip (1993, 2006) that perceptually based adaptations are pre-grammatical, in the sense that they are computed before the phonological grammar *per se* comes into play. Indeed, in our model, loanword adaptations are the reflex of a process that applies during phonetic decoding, whereas native alternations are the grammatical equivalent of the processes computed during phonological encoding. Loanword adaptations and native alternations are thus fundamentally different. Note, though, that they also share an important characteristic: they both consist of transformations that result in legal phonological surface forms (loanword adaptations in perception and native alternations in production). This accounts for the fact that if the same phonotactically illegal structure is present both in the source forms of loanwords and in the underlying forms of native words, the corresponding loanword adaptation and native alternation typically involve the same transformations. Indeed, there are not many ways to transform an illegal form into a legal one in an economical way; native alternations and loanword adaptations thus often go hand in hand. Our model leaves open, however, the possibility of conflicts between loanword adaptations and native alternations, that is, of cases in which a given illegal form is treated differently, depending upon whether it concerns an underlying form of a native word or the source form of a loanword. It turns out that such conflicts indeed occur.

---

different levels of phonological knowledge. The model presented here is not the processing equivalent of Boersma's model, though. That is, the two models are not isomorphic.

Illegal structure	Loanword adaptation	Native alternation
coda consonants in Maori (Yip 2002*)	*vowel epenthesis* [wuːru] < wool	*consonant deletion* /maːtur/ → [maːtu] 'know'
coda ɲ in Lama (Ulrich 1997*)	*vowel epenthesis* [fiɲə̥] < Fr. vigne [viɲ] 'vineyard'	*fronting*[19] /ti+ɲə̥/ → [tîːn] 'elephants'
coda release features in Korean (Kang 1996*, Kenstowicz & Sohn 2001, Kenstowicz & Suchato 2006*)	*vowel epenthesis* [kʰocʰi] < coach [pos'ɨ] < boss	*suppression* /nacʰ/ → [nat] 'face' /nas/ → [nat] 'sickle'
obstruent-nasal clusters in Korean (Kang 1996)	*vowel epenthesis*[20] [pʰikʰɨnik] < picnic	*nasalisation* /kuk+mul/ → [kuŋmul] 'soup' /top+nɨn/ → [tomnɨn] 'help-REL'
consonant clusters in Korean (Kenstowicz & Suchato 2006*)	*vowel epenthesis* [peltʰɨ] < belt	*consonant deletion* /talk/ → [tak] 'chicken'
obstruent-liquid clusters in Fula (Paradis 1992, Paradis & LaCharité 1997)	*medial epenthesis* [taːbal] < Fr. table [tabl] 'table'	*final epenthesis* /sokl+ka/ → [soklaka] 'need'
monomoraic final syllables in Thai (Kenstowicz & Suchato 2006*)	*gemination* [dāːtâːtaːtaː] < data	*glottal stop epenthesis* /pʰrá/ → [pʰráʔ] 'monk'
je-sequences in Japanese (Itô & Mester 1995)	*vocalisation*[21] [iemeN] < Yemen	*deletion* /moj+eru/ → [moeru] 'burn-INTRANS'
intervocalic stops in Malayalam (Mohanan & Mohanan 2003*)	*gemination*[22] [beːkːar] < baker	*voicing* /makan/ → [magan] 'son'

Table II

Conflicts between loanword adaptations and native alternations. References that explicitly discuss the existence of a conflict are marked with *.

Table II presents various examples; all loanwords are from English unless otherwise indicated.

[19] Word-final schwa is deleted after sonorants, with compensatory lengthening applying to the preceding vowel (cf. /miˌ-rə̥/ → [míːr] 'nose').

[20] This adaptation is found in older speakers only; younger speakers apply nasalisation, as in native words.

[21] Alternatively, certain loanwords undergo deletion, as found in the native vocabulary.

[22] Strictly speaking this is not a case of loanword adaptation: the process of gemination takes place in Malayee English, a variety of English used by speakers of Malayalam.

Clearly, in all of these cases perceptual minimality can be at stake in either the native alternation or the loanword adaptation, but not both. For one of them, it seems likely *a priori* that it is the loanword adaptation and not the native alternation that constitutes a perceptually minimal change: this is the Korean case of epenthesis in loanwords *vs.* suppression of release features in the native phonology. Given that word-final stops are strictly unreleased in Korean, the perceptual distance between, for instance, [c^{h}] and [c^{h}i] (the surface form in loanwords) is probably smaller than that between [c^{h}] and [t$^{\neg}$] (the surface form in native words). Ultimately, only phonetic measurements and/or perception experiments can shed light on the empirical question of which changes are minimal in cases of conflict.

Our psycholinguistic model should be compared to alternative, linguistic, models, in which loanword adaptations are computed by the phonological grammar. In particular, Kang (2003), proposes, following Steriade (2001), that the phonological grammar is supplied with correspondence constraints that demand perceptual similarity between input and output. Crucially, the input to the grammar can be either an underlying form (in the case of a native alternation) or a foreign surface form (in the case of a loanword adaptation). This model, which has been adopted by Adler (2006) and Shinohara (2006), thus accounts for perceptual minimality effects in both native alternations and loanword adaptations. The existence of conflicts between loanword adaptations and native alternations, though, is problematic, since it shows that the ranking of the constraint demanding perceptual similarity can depend on the type of input, loanword or native word. Of course, individual languages might have a whole family of constraints demanding perceptual similarity that are ranked differently. In the cases presented in Table II, however, only one such constraint is at play, since the loanword adaptations and the native alternations repair the same illegal structures.

An obvious solution to this problem consists of allowing constraints demanding perceptual similarity to apply in the case of loanword adaptations only. This solution appears to be adopted in Kenstowicz & Suchato's (2006) study of loanwords in Thai. Noting the conflict reported in Table II between gemination in loanwords and glottal stop epenthesis in native words, the authors propose a set of high-ranked output-to-output faithfulness constraints. These constraints exercise their influence in the computation of loanword adaptations, where candidate adapted forms are evaluated in light of a foreign surface form. However, they are simply irrelevant for the derivation of surface forms from underlying forms, i.e. for the computation of native alternations. Native alternations and perceptually based loanword adaptations are thus driven by the need to satisfy the same markedness constraints, but differ with regard to the relevant faithfulness constraints: whereas native alternations are subject to input-to-output faithfulness, loanword adaptations are subject to output-to-output faithfulness.

The fact that both loanword adaptations and native alternations yield phonotactically legal surface forms, although in different ways and possibly with different results, is central to both our psycholinguistic and Kenstowicz & Suchato's (2006) linguistic proposal. Might it be the case, then, that Kenstowicz & Suchato's model is the grammatical equivalent of our psycholinguistic model? We answer this question in the negative, for two reasons. First, output-to-output faithfulness constraints take as their input phonetic representations, which can be fairly detailed but are discrete. The phonetic decoder in our model, by contrast, takes as its input a continuous representation of formant transitions. Second, nothing in Optimality Theory demands that output-to-output faithfulness constraints be top-ranked, and, therefore, that loanword adaptations be influenced by perception. Individual languages might have grammars in which some output-to-output faithfulness constraints are ranked high and others low. We even expect there to be languages in which output-to-output faithfulness is uniformly ranked so low that it never plays a role. In other words, languages are predicted to differ widely with respect to the role of perception in loanword adaptations. By contrast, in our psycholinguistic model, phonetic decoding applies regardless of the listener's native language and regardless of the source word. Consequently, perceptual assimilation is expected to be massively responsible for loanword adaptations.

This brings us to two final questions we should raise. The first one concerns the existence of loanword adaptations that are *not* due to perceptual assimilation and their place in the present model. As mentioned before, based on psycholinguistic evidence concerning the perception of non-native sound structures, we consider the default origin of loanword adaptations to be perceptual assimilation. There are at least two types of adaptations, though, which have different origins. Both have been mentioned before. First, many loanword adaptations are based on written input and arguably influenced by the orthography of the source language (see Iverson 2005 for examples, and Vendelin & Peperkamp 2006 for experimental evidence). Orthography being a metalinguistic characteristic, an account of this type of adaptation has not yet been integrated into our model.[23] Second, the phonological grammar of the borrowing language can play a role in loanword adaptations, most notably in the case of overgeneralisations to a default pattern, where perceptual assimilation arguably does not apply (see note 6). These adaptations are thus computed by the phonological encoding module. We do not exclude the existence of other types of adaptations that are likewise due to the phonological grammar of the borrowing language. However, given the primacy of perception over production and the automatic nature of perceptual assimilation, phonology can only be argued to come into play once perceptual

[23] Of course, a great many loanwords enter the borrowing language via both oral and written input. Disentangling the respective roles of orthography and perception in those cases is an issue that largely remains to be investigated.

assimilation is shown either not to take place or to be followed by further adjustments.

The second question concerns the existence of non-adaptations. Although most research in loanword phonology focuses on patterns of adaptation, examples of non-adaptations – i.e. non-native structures that are not modified when appearing in a loanword – are not hard to come by. For instance, Haugen (1950) mentions the otherwise unattested word-final presence in English of [ʒ] in *rouge*, a loanword from French. Many other examples can be found in, for instance, Holden (1976), Bauer (1985), Danesi (1985), Kang (1996), Paradis & LaCharité (1997), Itô & Mester (1999, 2001), Sicoli (2000), Ussishkin & Wedel (2003), Broselow (2004b), Kenstowicz & Suchato (2006) and Yip (2006). The question as to how non-adaptations should be dealt with has received some attention, too. Specifically, both Holden (1976) and Itô & Mester (1999, 2001) propose that markedness constraints that are never violated in the native phonology are in fact not all equally strong. The weaker an individual markedness constraint, the greater the likelihood that it can be violated in loanwords, giving rise to non-adaptations. Although such an analysis can easily be implemented within, for instance, Optimality Theory, it seems rather ad hoc in that it leaves unanswered the question as to *why* some constraints are weaker than others. Indeed, the native phonology provides no evidence for such a hierarchy of strengths, all the constraints in question being obeyed in native words. In our model, non-adaptations can arise only if a given non-native structure is not perceptually assimilated. Some non-native sounds and sound structures are indeed easier to perceive than others (Best 1994), thus making non-adaptation become more likely (for discussion along these lines of a concrete example, see Crawford, in press). We therefore make the empirical prediction that there is a correlation between ease of perception, as established in discrimination experiments, and the occurrence of non-adaptations.[24]

6 Conclusion

In the present article, we have considered an asymmetry in the Japanese adaptation of word-final [n] in loanwords from English and French: whereas in the former, [n] is adapted as a moraic nasal consonant, in the

[24] Alternatively, Ussishkin & Wedel (2003) propose that the question of whether a non-native structure is adapted or not depends upon articulatory ease, which they define as a function of the native language phonology (rather than as an inherent property of speech sounds). We disagree with their proposal, because it presupposes that non-native structures are never perceptually assimilated. However, for those cases in which perceptual assimilation arguably is not at stake, we acknowledge that ease of articulation can play a role. An example is provided by Zulu clicks, which have been shown by Best *et al*. (1988) to by and large escape perceptual assimilation in English listeners (due to the fact that they are acoustically so distant from any available native phonetic category), but which are nonetheless not likely to enter the periphery of the English sound system by means of loanword non-adaptations. Indeed, English speakers have no experience with the (linguistic) articulation of click sounds; hence, they would adapt Zulu clicks for articulatory reasons.

latter it is adapted with a final epenthetic vowel. We have provided experimental evidence that this asymmetry is due to phonetic differences in the realisation of word-final [n] in English and French, and, consequently, to the way in which English and French word-final [n] are perceived by native speakers of Japanese. In particular, French but not English word-final [n] has a strong vocalic release that is assimilated to the Japanese vowel [ɯ] during phonetic decoding, a process that is beyond conscious awareness and applies in both monolinguals and bilinguals. We have also proposed a psycholinguistic model that accounts for both similarities and differences between loanword adaptations and native alternations.

To conclude, loanword adaptations are a real speech phenomenon, arising during contact between speakers of the borrowing language with speakers of the source language. Based on psycholinguistic evidence concerning the perception of non-native sound patterns, we consider all loanword adaptations that do not represent generalisations to a default pattern and that are not influenced by orthography to originate in perceptual assimilation, unless evidence to the contrary is provided. This does not mean that these loanword adaptations are completely *identical* to perceptual assimilations. Indeed, whereas the phenomenon of perceptual assimilation shows both intra- and inter-subject variation, loanword adaptations tend to be homogeneous across speakers. This suggests that they are subject to a process of standardisation within the population of speakers of the borrowing language. Moreover, the outcome of this standardisation may change over time, as shown by the fact that a given non-native structure may be adapted in older loanwords but not in newer ones (see e.g. Haugen 1953, Crawford, in press). Our working hypothesis, then, is that loanword adaptations reflect the average result of perceptual assimilation as found in most speakers (whether mono- or bilingual). In future research, we plan to focus on this process of standardisation, for instance by examining how loanwords spread through a population and how their phonological shape evolves as a function of time and space (Haugen 1953, Poplack *et al*. 1988, Sicoli 2000, Crawford, in press).

Appendix

Training items	*Test items*			
/sɔmi'gɑn/	/dɛbɑ'min/	/mɑbi'jun/	/pimɑ'gin/	/sɔgi'nɑn/
/dɑbi'tɛn/	/dɛbi'pin/	/mibɑ'pɑn/	/pɔmɑ'zɛn/	/tɛbɑ'nɔn/
/pɑni'bɔn/	/dɛpɑ'nun/	/mɔbi'gɔn/	/sɛgɑ'pɛn/	/tɛgi'dɔn/
/mipɑ'kun/	/dɑmi'bin/	/pɑbi'mun/	/sɛmi'tɑn/	/tɑpi'gɛn/
/tɛmi'kin/	/mɛbɑ'zɔn/	/pɑmi'jɑn/	/sɑbi'mɛn/	/tɔpɑ'sun/

Carrier sentences (Experiment 1 only)

English Yesterday I met __, Louise and Judy.

French Hier j'ai rencontré __, Laurence et Julie.

REFERENCES

Adler, Allison N. (2006). Faithfulness and perception in loanword adaptation: a case study from Hawaiian. *Lingua* **116**. 1024–1045.

Aoyama, Katsura, James Emil Flege, Susan G. Guion, Reiko Akahane-Yamada & Tsuneo Yamada (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *JPh* **32**. 233–250.

Bauer, R. (1985). The expanding syllabary of Hong Kong Cantonese. *Cahiers de Linguistique Asie Orientale* **14**. 99–111.

Best, Catherine T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In Judith C. Goodman & Howard C. Nusbaum (eds.) *The development of speech perception: the transition from speech sounds to spoken words*. Cambridge, Mass.: MIT Press. 167–224.

Best, Catherine T. & Winifred Strange (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *JPh* **20**. 305–330.

Best, Catherine T., Gerald W. McRoberts & Nomathemba M. Sithole (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* **14**. 345–360.

Boersma, Paul (1999). On the need for a separate perception grammar. Ms, University of Amsterdam. Available as ROA-358 from the Rutgers Optimality Archive.

Broselow, Ellen (2004a). Language contact phonology: richness of the stimulus, poverty of the base. *NELS* **34**. 1–21.

Broselow, Ellen (2004b). Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism* **8**. 51–65.

Crawford, Clifford (in press). The role of loanword diffusion in changing adaptation patterns: a study of coronal stops in Japanese borrowings. *Working Papers of the Cornell Phonetics Laboratory* **16**. 32–56.

Danesi, Marcel (1985). *Loanwords and phonological methodology*. Ville LaSalle: Didier.

Davidson, Lisa (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology* **24**. 261–286.

Dehaene-Lambertz, Ghislaine, Emmanuel Dupoux & Ariel Gout (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study. *Journal of Cognitive Neuroscience* **12**. 635–647.

Dell, François (1973). *E* muet: fiction graphique ou réalité linguistique? In Stephen R. Anderson & Paul Kiparsky (eds.) *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston. 26–50.

Dell, François (1995). Consonant clusters and phonological syllables in French. *Lingua* **95**. 5–26.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler (1999). Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* **25**. 1568–1578.

Dupoux, Emmanuel, Christophe Pallier, Núria Sebastián & Jacques Mehler (1997). A destressing 'deafness' in French? *Journal of Memory and Language* **36**. 406–421.

Dupoux, Emmanuel, Sharon Peperkamp & Núria Sebastián-Gallés (2001). A robust method to study stress 'deafness'. *JASA* **110**. 1606–1618.

Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete & Sharon Peperkamp (2008). Persistent stress 'deafness': the case of French learners of Spanish. *Cognition* **106**. 682–706.

Flege, James Emil, Naoyuki Takagi & Virginia Mann (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *JASA* **99**. 1161–1173.

Gandour, Jack (1983). Tone perception in Far Eastern languages. *JPh* **11**. 149–175.

Gbéto, Flavien (2000). *Les emprunts linguistiques d'origine européenne en Fon* (*Nouveau Kwa, Gbe : Bénin*). Cologne: Rüdiger Köppe.
Goad, Heather & Kathleen Brannen (2003). Phonetic evidence for phonological structure in syllabification. In Jeroen van de Weijer, Vincent J. van Heuven & Harry van der Hulst (eds.) *The phonological spectrum*. Vol. 2: *Suprasegmental structure*. Amsterdam & Philadelphia: Benjamins. 3–30.
Golston, Chris & Phong Yang (2001). White Hmong loanword phonology. In Caroline Féry, Antony Dubach Green & Ruben van de Vijver (ed.) *Proceedings of HILP 5*. Potsdam: University of Potsdam. 40–57.
Goto, Hiromu (1971). Auditory perception by normal Japanese adults of the sounds 'L' and 'R'. *Neuropsychologia* **9**. 317–323.
Hallé, Pierre A., Juan Segui, Uli Frauenfelder & Christine Meunier (1998). Processing of illegal consonant clusters: a case of perceptual assimilation? *Journal of Experimental Psychology : Human Perception and Performance* **24**. 592–608.
Harris, John (1997). Licensing Inheritance: an integrated theory of neutralisation. *Phonology* **14**. 315–370.
Haugen, Einar (1950). The analysis of linguistic borrowing. *Lg* **26**. 210–231.
Haugen, Einar (1953). *The Norwegian language in America : a study in bilingual behavior*. Vol. 2: *The American dialects of Norwegian*. Philadelphia: University of Pennsylvania Press.
Holden, Kyril (1976). Assimilation rates of borrowings and phonological productivity. *Lg* **52**. 131–147.
Hyman, Larry M. (1970). The role of borrowing in the justification of phonological grammars. *Studies in African Linguistics* **1**. 1–48.
Ito, Chiyuki, Yoonjung Kang & Michael Kenstowicz (2006). The adaptation of Japanese loanwords into Korean. In Feng-fan Hsieh & Michael Kenstowicz (eds.) *Studies in loanword phonology* (*MITWPL* **52**). 65–104.
Itô, Junko & Armin Mester (1995). Japanese phonology. In John A. Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell. 817–838.
Itô, Junko & Armin Mester (1999). The phonological lexicon. In Tsujimura (1999). 62–100.
Itô, Junko & Armin Mester (2001). Covert generalizations in Optimality Theory: the role of stratal faithfulness constraints. *Studies in Phonetics, Phonology, and Morphology* **7**. 273–299.
Iverson, Gregory K. (2005). The principling role of Korean in phonological adaptation. In Sang-Oak Lee (ed.) *The 30th anniversary publication of the International Circle of Korean Linguistics*. Seoul: Pagijong. 141–159.
Iverson, Gregory K. & Ahrong Lee (2006). Perception of contrast in Korean loanword adaptation. *Korean Linguistics* **13**. 49–87.
Jacobs, Haike & Carlos Gussenhoven (2000). Loan phonology: perception, salience, the lexicon and OT. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.) *Optimality Theory : phonology, syntax, and acquisition*. Oxford: Oxford University Press. 193–210.
Jacquemot, Charlotte, Christophe Pallier, Denis LeBihan, Stanislas Dehaene & Emmanuel Dupoux (2003). Phonological grammar shapes the auditory cortex: a functional Magnetic Resonance Imaging study. *Journal of Neuroscience* **23**. 9541–9546.
Kang, Hyunsook (1996). English loanwords in Korean. *Studies in Phonetics, Phonology, and Morphology* **2**. Seoul: Hankwukmunhwasa. 21–47.
Kang, Yoonjung (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* **20**. 219–273.
Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud (1990). Constituent structure and government in phonology. *Phonology* **7**. 193–231.

Kenstowicz, Michael (2004). Review article: the role of perception in loanword phonology. (Review of Gbéto 2000.) *Studies in African Linguistics* **32**. 95–112.
Kenstowicz, Michael & Hyang-Sook Sohn (2001). Accentual adaptation in North Kyungsang Korean. In Michael Kenstowicz (ed.) *Ken Hale : a life in language*. Cambridge, Mass: MIT Press. 239–270.
Kenstowicz, Michael & Atiwong Suchato (2006). Issues in loanword adaptation: a case study from Thai. *Lingua* **116**. 921–949.
Kenstowicz, Michael & Christian Uffmann (eds.) (2006). *Loanword phonology : current issues*. Special issue. *Lingua* **116**. 921–1194.
Kim, Soohee & Emily Curtis (2002). Phonetic duration of English /s/ and its borrowing in Korean. In Noriko M. Akatsuka & Susan Strauss (eds.) *Japanese/ Korean Linguistics 10*. Palo Alto: CSLI. 406–419.
Kiriloff, C. (1969). On the auditory perception of tones in Mandarin. *Phonetica* **20**. 63–67.
Kondo, Mariko (1997). *Mechanisms of vowel devoicing in Japanese*. PhD dissertation, University of Edinburgh.
Kubozono, Haruo (1999). Mora and syllable. In Tsujimura (1999). 31–61.
Kuhl, Patricia K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences* **97**. 11850–11857.
LaCharité, Darlene & Carole Paradis (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *LI* **36**. 223–258.
Lahiri, Aditi & William Marslen-Wilson (1991). The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition* **38**. 245–294.
Lee, Lisa & Howard C. Nusbaum (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception and Psychophysics* **53**. 157–165.
Levelt, Willem J. M. (1989). *Speaking : from intention to articulation*. Cambridge, Mass.: MIT Press.
Levelt, Willem J. M., Ardi Roelofs & Antje S. Meyer (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* **22**. 1–38.
Lively, Scott E., John S. Logan & David B. Pisoni (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *JASA* **94**. 1242–1255.
Lovins, Julie (1975). *Loanwords and the phonological structure of Japanese*. Bloomington: Indiana University Linguistics Club.
Massaro, Dominic W. & Michael M. Cohen (1983). Phonological context in speech perception. *Perception and Psychophysics* **34**. 338–348.
Mielke, Jeff (2005). Modeling distinctive feature emergence. *WCCFL* **24**. 281–289.
Mohanan, Tara & K. P. Mohanan (2003). Towards a theory of constraints in OT: emergence of the not-so-unmarked in Malayalee English. Available as ROA-601 from the Rutgers Optimality Archive.
Nakamura, Kimihiro & Emmanuel Dupoux (in preparation). *Perceptual epenthesis in Japanese-French bilinguals*.
Paradis, Carole (1992). *Lexical phonology and morphology : the nominal classes in Fula*. New York: Garland.
Paradis, Carole & Darlene LaCharité (1997). Preservation and minimality in loanword adaptation. *JL* **33**. 379–430.
Peperkamp, Sharon (2005). A psycholinguistic theory of loanword adaptations. *BLS* **30**. 341–352.
Peperkamp, Sharon & Emmanuel Dupoux (2003). Reinterpreting loanword adaptations: the role of perception. In M. J. Solé, D. Recasens & J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Causal Productions. 367–370.

Poplack, Shana, David Sankoff & Christopher Miller (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* **26**. 47–104.
Rabiner, Lawrence & Biing-Hwang Juang (1993). *Fundamentals of speech recognition*. Englewood Cliffs: Prentice Hall.
Rose, Yvan (1999). A structural account of root node deletion in loanword phonology. *Canadian Journal of Linguistics* **44**. 359–404.
Shibatani, Masayoshi (1990). Japanese. In Bernard Comrie (ed.) *The world's major languages*. London & Sydney: Croom Helm. 855–880.
Shinohara, Shigeko (1997). *Analyse phonologique de l'adaptation japonaise de mots étrangers*. PhD dissertation, Université de Paris III.
Shinohara, Shigeko (2000). Default accentuation and foot structure in Japanese: evidence from Japanese adaptations of French words. *Journal of East Asian Linguistics* **9**. 55–96.
Shinohara, Shigeko (2006). Perceptual effects in final cluster reduction patterns. *Lingua* **116**. 1046–1078.
Sicoli, Mark (2000). Loanwords and contact-induced phonological change in Lachixío Zapotec. *BLS* **25**. 395–406.
Silverman, Daniel (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* **9**. 289–328.
Smith, Jennifer (2006). Loan phonology is not all perception: evidence from Japanese loan doublets. In Timothy J. Vance & Kimberly A. Jones (eds.) *Japanese/Korean Linguistics 14*. Palo Alto: CLSI. 63–74.
Steriade, Donca (2001). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Ms, University of California, Los Angeles.
Takagi, Naoyuki & Virginia Mann (1994). A perceptual basis for the systematic phonological correspondences between Japanese loan words and their English source words. *JPh* **22**. 343–356.
Tranel, Bernard (1987). *The sounds of French : an introduction*. Cambridge: Cambridge University Press.
Tsujimura, Natsuko (ed.) (1999). *The handbook of Japanese linguistics*. Malden, Mass. & Oxford: Blackwell.
Ulrich, Charles (1997). Loanword adaptation in Lama: testing the TCRS model. *Canadian Journal of Linguistics* **42**. 415–463.
Ussishkin, Adam & Andrew Wedel (2003). Gestural motor programs and the nature of phonotactic restrictions: evidence from loanword phonology. *WCCFL* **22**. 505–518.
Vendelin, Inga & Sharon Peperkamp (2004). Evidence for phonetic adaptation of loanwords: an experimental study. *Actes des Journées d'Etudes Linguistiques de l'Université de Nantes*. 127–131.
Vendelin, Inga & Sharon Peperkamp (2006). The influence of orthography on loanword adaptations. *Lingua* **116**. 996–1007.
Werker, Janet F. & Richard C. Tees (1984). Phonemic and phonetic factors in adult cross-language speech perception. *JASA* **75**. 1866–1878.
Yip, Moira (1993). Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* **2**. 261–291.
Yip, Moira (2002). Necessary but not sufficient: perceptual influences in loanword phonology. *Journal of the Phonetic Society of Japan* **6:1**. 4–21.
Yip, Moira (2006). The symbiosis between perception and grammar in loanword phonology. *Lingua* **116**. 950–975.